

RESEARCH ARTICLE

Using balanced acceptance sampling as a master sample for environmental surveys

Paul van Dam-Bates¹  | Oliver Gansell² | Blair Robertson³¹Planning Monitoring and Reporting,
Department of Conservation, Christchurch,
New Zealand²Planning Monitoring and Reporting,
Department of Conservation, Hamilton,
New Zealand³School of Mathematics and Statistics,
University of Canterbury, Christchurch,
New Zealand**Correspondence**

Paul van Dam-Bates

Email: paul.vandambates@gmail.com

Handling Editor: Robert Freckleton

Abstract

1. Well-designed environmental monitoring programmes for management organisations are important for evidence-based decision making. However, many environmental problems are not single agency issues that require intervention or monitoring at one spatial scale. A master sample can be used to coordinate and scale monitoring designs to ensure consistency in information gathered and robustness of estimators at the different spatial scales.
2. We propose using balanced acceptance sampling (BAS) to generate a master sample. In this context, practical applications and justification of BAS as a master sample are addressed. These include sample generation, stratification, unequal probability sampling, rotating panel designs, and regional intensification. A method for incorporating legacy sites is also provided.
3. Using BAS as a master sample is conceptually simple, gives good spatial balance over different spatial scales, and is computationally efficient to generate. An example for terrestrial biodiversity monitoring in New Zealand is provided.
4. Environmental monitoring can benefit from increased coordination between agencies. A master sample is an excellent way to incorporate coordination directly into the sample design. BAS improves on methods previously described and provides an effective method to monitor populations at multiple spatial scales.

KEYWORDS

environmental monitoring, legacy monitoring, master sample, spatial balance

1 | INTRODUCTION

Environmental management agencies rely on the results of monitoring to answer questions about the success of their policies and programmes. Monitoring is often designed to address information needs for a particular site or small set of sites. If the study is poorly designed, it can fail to provide meaningful data to inform management and policy decision making (Field, O'Connor, Tyre, & Possingham, 2007; Legg & Nagy, 2006; Nichols & Williams, 2006). Extrapolating from these studies to answer larger scale questions can bias the

estimates as single sites are rarely representative of a broader region (Dixon, Olsen, & Kahn, 1998; Peterson, Urquhart, & Welch, 1999).

Increasingly, monitoring on a large scale is needed to inform management needs and assess progress towards targets concerned with global biodiversity change (Buckland, Magurran, Green, & Fewster, 2005; Magurran et al., 2010; Noss, 1999; Pereira & Cooper, 2006). Monitoring objectives and sample areas of national, regional, and local agencies often overlap creating efficiencies if the different groups coordinate their effort. Coordinating requires consistent formulation of goals and objectives, selection of indicators and measures, field

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society

protocols and sample design (Larsen, Olsen, & Stevens, 2008; Fancy, Gross, & Carter, 2009; Reynolds, Knutson, Newman, Silverman, & Thompson, 2016). If one agency establishes monitoring locations using standard methods and sample design, others can use that data for their own purposes, reducing the need to establish more monitoring. By agencies working together and through a well set out design process, as described in Reynolds et al. (2016), the chances of monitoring being successful are higher and concerns about extrapolating estimates from disparate data sources are also reduced.

One way to coordinate sample design is to develop a master sample; a set of points that can be subsampled for different monitoring activities. This was first proposed by King (1945), but only recently has been introduced to environmental monitoring (Larsen et al., 2008; Theobald, 2016) with implementation in the Pacific Northwest of the United States. Different studies drawing samples from the master sample enhances collaboration within and between agencies to reduce duplication of effort. Additionally, consistent sample design has benefits when making estimates using data from multiple sources. Similar to providing standard field methods, the master sample provides standardised locations for sampling that ensures objective, unbiased estimation of the population parameters of interest. The coordinating body requires the user to define the objectives and sample frame clearly before gaining access to the sample points.

The sampling method chosen should be flexible enough for a variety of users and study designs to be effective for coordination. Monitoring can take place on different spatial scales such as a national monitoring programme or a local one investigating the impact of management action. When designing an individual study, identifying heterogeneity and using stratification (Yoccoz, Nichols, & Boulmier, 2001) or unequal probability sampling (Stevens, 1997) can produce more precise estimates. The study may need a unique balance of status and trend estimation which can be done by defining panels that have different revisit schema (McDonald, 2003; Skalski, 1990; Stevens & Olsen, 1999). In all these cases, the subsamples used should be unbiased and representative.

There are many ways to generate effective samples which could be used to coordinate monitoring. A simple random sample is unbiased but is less efficient than spatially balanced designs in the presence of spatial autocorrelation (Cochran, 1946; Grafström & Lundström, 2013). A design is spatially balanced if the sample is well-spread over the population – a sample with few clumps and voids. A systematic sample can be considered near perfect spatial balance but is less flexible to changes in sample size making it a poor choice. Stevens and Olsen (2004) introduced generalised random tessellation stratified (GRTS) design, a spatially balanced design that is frequently used in environmental monitoring (Collier & Olsen, 2013; Fancy et al., 2009; Thompson, Miller, Mortenson, & Woodward, 2011). Generalised random tessellation stratified hierarchically orders a population using a base four numbering scheme and then selects a systematic sample from the ordered population. There is also the Local Pivotal Method (LPM) (Grafström, Lundström, & Schelin, 2012). LPM iteratively updates each sampling unit's inclusion probability in a way that makes it very unlikely to include neighbouring units in a sample. Once n units have an inclusion probability of one, the sample

is released. Although the spatial balance of LPM is better than GRTS, it is computationally prohibitive on large populations. For large populations, Grafström, Saarela and Ene (2014) introduced a new rapid implementation of LPM, called suboptimal LPM. LPM has better spatial balance, but suboptimal LPM is computationally feasible on large populations. Another spatially balanced design is balanced acceptance sampling (BAS) (Robertson, Brown, McDonald, & Jaksons, 2013). It uses a quasi-random number sequence to generate spatially balanced points. Similar to GRTS, the outcome of the sequence is an ordered set of points such that any contiguous subsample maintains spatial balance.

Generalised random tessellation stratified design has been used to generate environmental monitoring master samples (Larsen et al., 2008). The design is particularly useful for generating master samples because GRTS points are ordered using a reverse hierarchical ordering strategy that ensures that all contiguous subsamples are also spatially balanced (Stevens & Olsen, 2004). By taking a large GRTS oversample, an ordered master sample can be obtained from which spatially balanced subsamples can be drawn. However, once an oversample is chosen, it is not possible to generate additional points and this needs to be accounted for at the planning stage. Theobald (2016) also uses an adaptation of GRTS, Reversed Randomised Quadrant-Recursive Raster (RRQRR), implemented in ArcGIS software (Theobald et al., 2007) to coordinate monitoring effort. The authors' are not aware of an ordering strategy for the LPM methods and hence, it is not clear how these methods could be used for oversampling. BAS, similar to GRTS, creates an ordered set of points such that any contiguous subsample maintains spatial balance. To generate a master sample with BAS, a random-start is chosen and after that an infinite set of points exist for the sample. Hence, the oversample size does not need to be specified.

The purpose of this paper is to develop a master sample for environmental monitoring with a focus on terrestrial sampling of an area frame in which all subregions have positive area. We investigate using BAS to generate a master sample; how the points will be generated and then used in a wide variety of ways. These include adapting to different spatial scales, stratification and unequal probability sampling, changes in boundaries or resources, revisitation structure (panel design) and how to include legacy monitoring programmes. We then provide an example for how this could be applied to coordinate biodiversity monitoring at the regional and national level in New Zealand.

2 | MATERIALS AND METHODS

2.1 | Point selection

Two-dimensional BAS points are drawn from a random-start Halton sequence $\{\mathbf{x}_k\}_{k=1}^{\infty} \subset [0,1]^2$. The i th coordinate of each point in the sequence has an associated base b_i , with $b_1 = 2$ and $b_2 = 3$. The i th coordinate of the k th point in this sequence is (Robertson, McDonald, Price, & Brown, 2017)

$$x_k^{(i)} = \sum_{j=0}^{\infty} \left\{ \left\lfloor \frac{u_i + k}{b_i^j} \right\rfloor \bmod b_i \right\} \frac{1}{b_i^{j+1}},$$

where u_i is a random non-negative integer and $\lfloor x \rfloor$ is the floor function — the largest integer that is less than or equal to x . For example, the first coordinate of the second point with $u_1 = 1$ and $b_1 = 2$ is

$$\begin{aligned} x_2^{(1)} &= \left(\left\lfloor \frac{3}{1} \right\rfloor \bmod 2 \right) \frac{1}{2} + \left(\left\lfloor \frac{3}{2} \right\rfloor \bmod 2 \right) \frac{1}{4} + \left(\left\lfloor \frac{3}{4} \right\rfloor \bmod 2 \right) \frac{1}{8} \\ &= \frac{1}{2} + \frac{1}{4} + 0 \\ &= \frac{3}{4}. \end{aligned}$$

The two-dimensional random-start Halton sequence is

$$\{\mathbf{x}_k\}_{k=1}^{\infty} = \left\{ x_k^{(1)}, x_k^{(2)} \right\}_{k=1}^{\infty}. \quad (1)$$

Setting $u_1 = u_2 = 0$ gives the classical two-dimensional Halton sequence (Halton, 1960).

The points from Equation 1 are then scaled to a minimal bounding box enclosing the study area. If \mathbf{x}_1 is not in the study area, new sequences are considered until one with \mathbf{x}_1 in the study area is found (Robertson et al., 2017). Starting from \mathbf{x}_1 , the first n scaled points in the study area define the BAS sample (Robertson et al., 2017). The BAS points are kept in the same order as they appear in Equation 1 and will have good spatial spread over the study area. Furthermore, any contiguous subset of the BAS sample will also have good spatial spread (Robertson et al., 2017).

The random integer vector in the sequence $\mathbf{u} = (u_1, u_2)$ is chosen so that $0 \leq u_i \leq 10^7$. This gives $\approx \lambda 10^{14}$ possible BAS samples of size n , where λ is the fraction of the bounding box occupied by the study area. By ensuring the random start comes from a large set of integers, the BAS points are uniformly distributed (Robertson et al., 2013). Once the random-start is selected an infinite number of BAS points exist over the study which constitutes the master sample. Higher dimensional points can be defined by using different co-prime bases for each additional dimension (e.g. $b_3 = 5$ when sampling from a $[0,1]^3$).

2.2 | Spatial scales

The master sample should work at different spatial scales to address national, regional, and local objectives. Let A be a measurable subset of the study area for which the master sample is defined. Because the master sample $\{\mathbf{x}_k\}_{k=1}^{\infty}$ is uniformly distributed over $[0,1]^2$ (Wang & Hickernell, 2000) and A is measurable, there exists a subsequence $\{\mathbf{z}_j\}_{j=1}^{\infty} \subset \{\mathbf{x}_k\}_{k=1}^{\infty}$ such that each $\mathbf{z}_j \in A$. Furthermore, $\{\mathbf{z}_j\}_{j=1}^n$ is a BAS sample of size n drawn from A , with its random start and bounding box defined by the master sample. Hence, BAS samples can be drawn from the master sample at any spatial scale within the study area of the master sample. This also means that a national sample can share points with monitoring at the local level (see Section 2.4).

2.3 | Stratification and unequal probability

Stratification with the master sample is essentially the same as taking a subsample for a specific measurable subset of the study area as described above. The i th stratum (measurable) has a subsequence $\{\mathbf{z}_j\}_{j=1}^{\infty} \subset \{\mathbf{x}_k\}_{k=1}^{\infty}$ such that each \mathbf{z}_j is in the stratum. The BAS sample

for the i th stratum is $\{\mathbf{z}_j\}_{j=1}^{n_i}$, where n_i is the sample size required. Hence, each stratum has its own BAS sample with its random start and bounding box defined by the master sample.

If unequal probability sampling is required, a third dimension is added to the bounding box. This extra dimension allows BAS to sample from an arbitrary inclusion density function $\pi(\mathbf{x})$ using an acceptance/rejection sampling strategy (Robertson et al., 2013). Specifically, a point $\mathbf{x}_k = (x_k^{(1)}, x_k^{(2)}, x_k^{(3)})$ is accepted if $(x_k^{(1)}, x_k^{(2)})$ is in the study area and $\pi(x_k^{(1)}, x_k^{(2)}) \leq \alpha x_k^{(3)}$, where α is the maximum value of $\pi(\mathbf{x})$. The impact of this is that some of the master sample points in Equation 1 will be skipped. Skipping points in Equation 1 changes the distribution of points in each BAS sample, with fewer points being drawn from areas where $\pi(\mathbf{x})$ is low.

2.4 | Changing boundaries and resources

For long-term monitoring programmes, the boundaries of study regions may change over time. This is easy to accommodate with the master sample, provided the changes are within the initial bounding box. Let A be a measurable study area whose boundaries changed, defining a new measurable study area B with $A \cap B \neq \emptyset$. If there are no sampled BAS points in $A \cap B$, points from the master sample are drawn to sample B . Otherwise, let \mathbf{x}_k be the sampled point in $A \cap B$ with the largest index k . A BAS sample in B , that includes sampled points from $A \cap B$, is achieved if all the master sample points that fall in B with indices less than k are sampled. If a smaller sample is desired in B , potentially due to a change in resources, then points with the larger indices in B are removed (See Figure 1). If more points are required, then points can be added from the master sample in B until the new sample size is achieved. Ensuring that BAS samples are drawn from each study region means that spatial balance and good sampling properties are maintained.

2.5 | Panel design

In environmental surveys that are repeated through time, some samples may be visited more frequently than others. Estimates of status and/or trend can be improved by balancing the number of new points sampled each year with repeated sampling on existing points (Urquhart & Kincaid, 1999). A panel is defined as all samples that have the same visitation schedule. The allocation of points within a panel as well as between panels should be representative and unbiased. A panel design is achieved using the master sample by choosing the subset of points $\{\mathbf{z}_j\}_{j=1}^{\infty}$ that fall within the study area and the order of the points define each panel. For panel 1, we use $\{\mathbf{z}_j\}_{j=1}^{n_1}$, for panel 2 we use $\{\mathbf{z}_j\}_{j=n_1+1}^{n_1+n_2}$ and so on, where n_i is the sample size for the i th panel. Defining panels in this way ensures that both the overall sample and each panel's sample are BAS designs. If additional points are needed after a full rotation, they are taken from the unsampled points in the master sample in the order that they appear. Note, when this occurs each panel may not be a true BAS sample but as shown in the legacy monitoring simulation below, adding BAS points to an existing sample does not significantly impact estimation and the sample

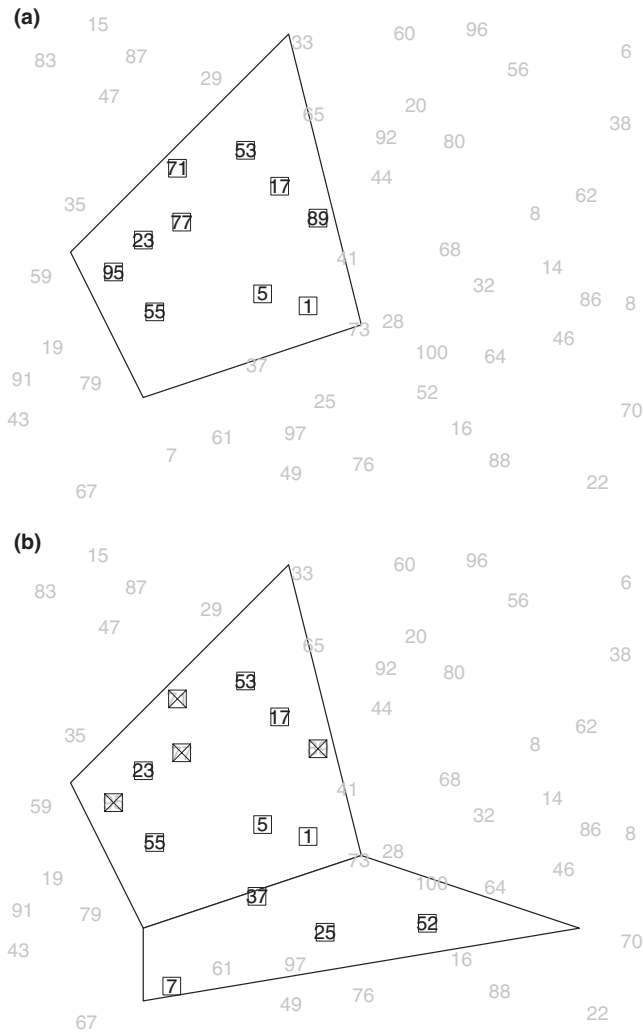


FIGURE 1 An example of changing boundaries using a BAS the master sample where each number denotes the index of the BAS point in the master sample. (a) shows 10 points drawn from the master sample in the study region. (b) shows an enlarged region and the 10 points drawn from the master sample in the region. Points 71, 77, 89 and 95 are removed and 7, 25, 37 and 52 are included in the new region. If the sample size is increased to 14, then points 61, 71, 77 and 89 would be included

will still be equi-probable and give unbiased estimators. If budgets change, points should be removed by last in, first out. Table 1 shows an example panel design.

TABLE 1 Example of a panel design in which panel 1 is sampled annually and panels 2–4 are sampled with a 2 year break in between described as $[1-0, (1-2)^3]$ in McDonald (2003). The sample size n and the points from the master sample in the study area $\{z_j\}$ are shown below, where an X indicates that the panel is sampled on that occasion

Panel	n	Sample	Sample occasion									
			1	2	3	4	5	6	7	8	9	10
1	20	$\{z_j\}_{j=1}^{20}$	X	X	X	X	X	X	X	X	X	X
2	10	$\{z_j\}_{j=21}^{30}$	X			X			X			X
3	10	$\{z_j\}_{j=31}^{40}$		X			X			X		
4	10	$\{z_j\}_{j=41}^{50}$			X			X			X	

2.6 | Incorporating legacy monitoring

A master sample is intended for coordinating large-scale monitoring. Often, there is legacy monitoring that may already be well designed and this should be accommodated. We will consider two different types of legacy monitoring designs: simple random sampling (SRS) and random-start systematic sampling (SS). These are equi-probable designs, where each sampling unit has an equal chance of being included in a sample. If the existing monitoring is insufficient, the BAS master sample can be used to draw additional units from the area. We consider two types of sample augmentation. In both cases, we assume n_l legacy units were originally sampled with equal inclusion probabilities from a population of N units, and that these will be augmented by n_b additional units. The first augmentation method we consider removes legacy units, draws a regular BAS sample from the remainder, and includes both legacy and augmented units in the sample. In this case, we assume that the total sample arose jointly from a single sampling operation and hence, the inclusion probabilities for the legacy and augmented units are equal, $\pi_i = (n_l + n_b)/N$.

The second augmentation method we consider uses the unequal probability BAS sampling method described above to down-weight units near legacy units (Foster et al., 2017). The method we employ increases the probability of inclusion smoothly with increasing distance D from legacy units. Following from Foster et al. (2017), we sample proportional to a Gaussian Kernel, $f(D) = 1 - \exp(-D^2/\sigma^2)$, where σ controls the area of influence around legacy units. It is necessary in this case to re-scale the inclusion probabilities so that they sum to n . We will call this second method of sample augmentation that utilises altered inclusion probabilities aBAS.

Spatially balanced designs use a local neighbourhood variance (LNV) estimator (Stevens & Olsen, 2003), but as the proportion of legacy units is increased the LNV tends to underestimate true variance (Foster et al., 2017). If the n_l legacy units and the n_b augmented units are considered separately, the population mean can be estimated using the convex combination

$$\hat{\mu} = \left(\frac{n_l}{n}\right) \hat{\mu}_l + \left(\frac{n_b}{n}\right) \hat{\mu}_b, \quad (2)$$

where $\hat{\mu}_l$ and $\hat{\mu}_b$ the sample means of the legacy and augmented units, respectively. The variance can be estimated as

$$\widehat{\text{var}}(\hat{\mu}) = \left(\frac{n_l}{n}\right)^2 \times \widehat{\text{var}}(\hat{\mu}_l) + \left(\frac{n_b}{n}\right)^2 \times \widehat{\text{var}}(\hat{\mu}_b), \quad (3)$$

where $\widehat{\text{var}}(\hat{\mu}_l)$ is the Sen, Yates and Grundy variance estimate using the legacy units and $\widehat{\text{var}}(\hat{\mu}_b)$ is the LNV variance estimate using spatially balanced units.

A simulation study was carried out to investigate the different methods to incorporate legacy monitoring. The sampling frame was defined as a 100×100 raster in $[0,1]^2$ and an estimate of the population mean and standard error was sought. The response value for each raster cell was defined as the integral of $f(x)$ over the cell. Three different functions were considered to define the population, a strong spatial trend (Robertson et al., 2013;

Grafström et al., 2012), the Peak function, and the Bird function. These functions are given in the online supplementary material section. Scenarios similar to Foster et al. (2017) using the program R (R Core Team, 2015) were run. We used an overall sample size of $n = 60$ and a number of legacy units ($n_l \in 3, 4, \dots, 57$) were generated either as SRS or SS. More units ($n_b = 60 - n_l$) were generated using GRTS (Kincaid & Olsen, 2016), BAS, aBAS (Foster, 2016) and SRS to achieve the full sample size of $n = 60$. Each scenario was run 1,000 times. When estimating the standard error of SS legacy units, we used SRS estimators, which provides conservative estimates in the presence of spatial autocorrelation (Aune-Lundberg & Strand, 2014; Strand, 2017). A detailed description of

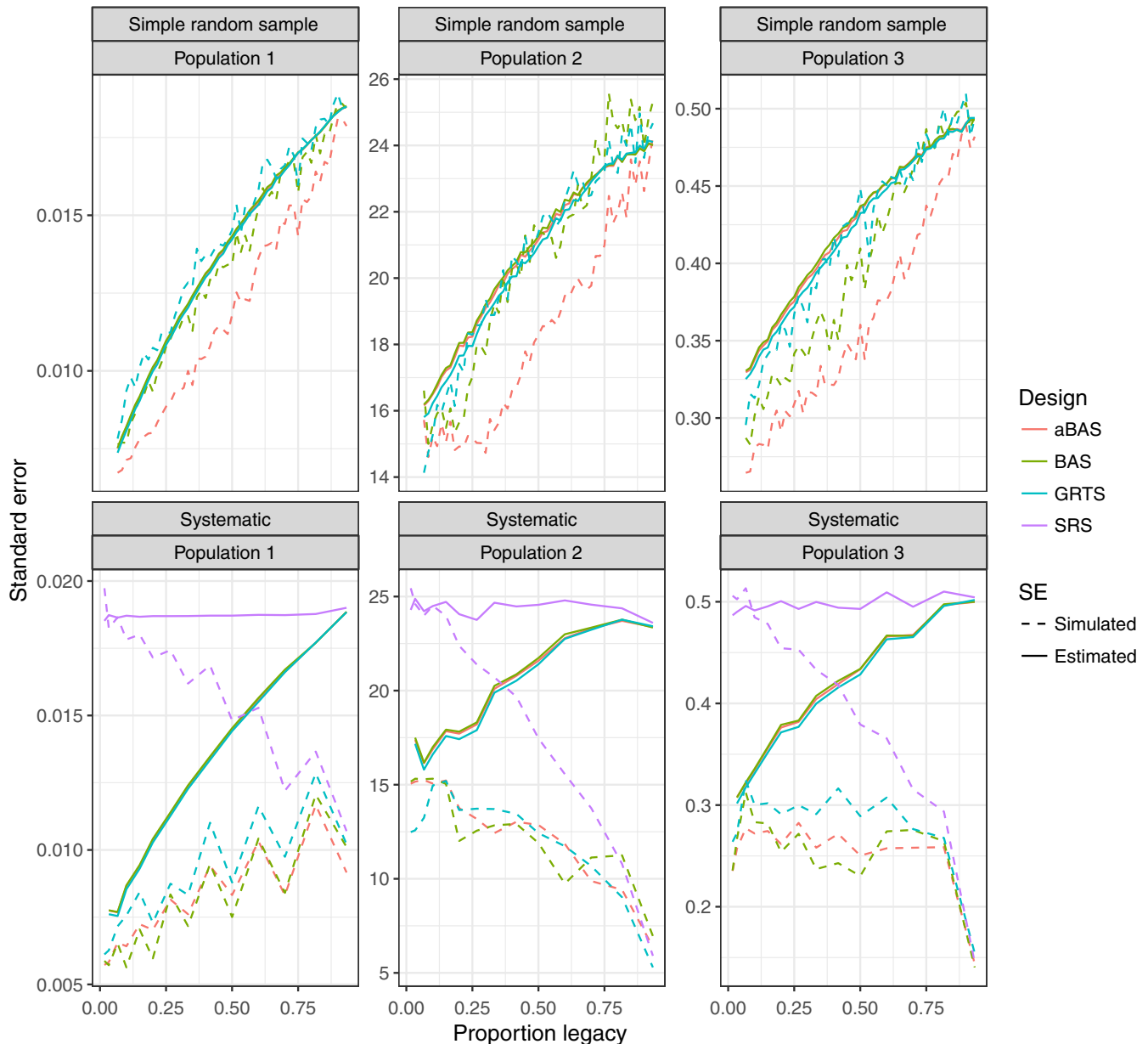


FIGURE 2 Results from the simulation study testing the impact of adding new samples from altered balanced acceptance sampling (aBAS), balanced acceptance sampling (BAS), generalised random tessellation stratified (GRTS), and simple random sampling (SRS) to existing legacy monitoring units were tested (see supplementary material section). The estimated (Equation 3) and simulated standard errors are shown

the simulation and functions used can be accessed in the supplementary material.

The results of the simulation can be seen in Figure 2. In all cases, adding some spread to the points improved precision over using SRS. aBAS performed the best when SRS legacy units were used, but had similar performances to BAS and GRTS for SS. The SS legacy units already had good spread and aBAS was not necessary to force better overall spatial balance. Population 3 has periodic structure, which made systematic sampling perform poorly because the spread of the legacy units matched the structure of the population. The standard error estimate from Equation 3 worked well for SRS, although conservative for aBAS, as well as SS. The legacy monitoring was improved by augmenting the samples with spatially balanced points. Assuming the legacy monitoring is an unbiased, representative and probabilistic sample, there are no issues with incorporating legacy samples with the master sample.

The aBAS design can work well for single studies or simple repeated studies when legacy monitoring is SRS or poorly spread. However, inflated variance is possible because the altered inclusion probabilities may not be positively correlated with the response variable (Robertson et al., 2017). In general, we recommend drawing spatially balanced points independent of the legacy monitoring. For design-based estimation, the variance estimate in Equation 3 should be used. This corrects the underestimation of LNV but requires there are four or more spatially balanced points as suggested in the *SPSURVEY* package in R (Kincaid & Olsen, 2016) and three or more legacy points. Otherwise, use LNV ($n_l < 3$) or the legacy variance estimator ($n_b < 4$) for all points. In practice, model-based estimation is often used and any analyses should account for the spatial aspect of the design as in Foster et al. (2017).

3 | APPLICATION: NEW ZEALAND TERRESTRIAL MONITORING

The New Zealand (NZ) Department of Conservation (DOC) is the lead biodiversity management agency in NZ, responsible for managing $\approx 30\%$ of NZ as public conservation land (PCL). Development of a national monitoring system has exposed the challenges in coordinating monitoring design to provide results meaningful at a local, regional and national scale. Increasingly partner environmental agencies (local government etc.) and central government expect cross-agency collaboration and coordination of systems and processes. Considerable effort has gone into a coordinated approach for indicators and measures and field protocols (Department of Conservation, 2016). There is an existing national sample of PCL, known as the National Level Monitoring (NLM) programme which is an 8-km systematic grid of $\approx 1,400$ monitoring sites (Coomes, Allen, Scott, Goulding, & Beets, 2002). The NLM programme focusses on status and trend monitoring at the national scale for key indicators of ecosystem structure and composition. Through collaboration with local government agencies, the grid has been extended off PCL.

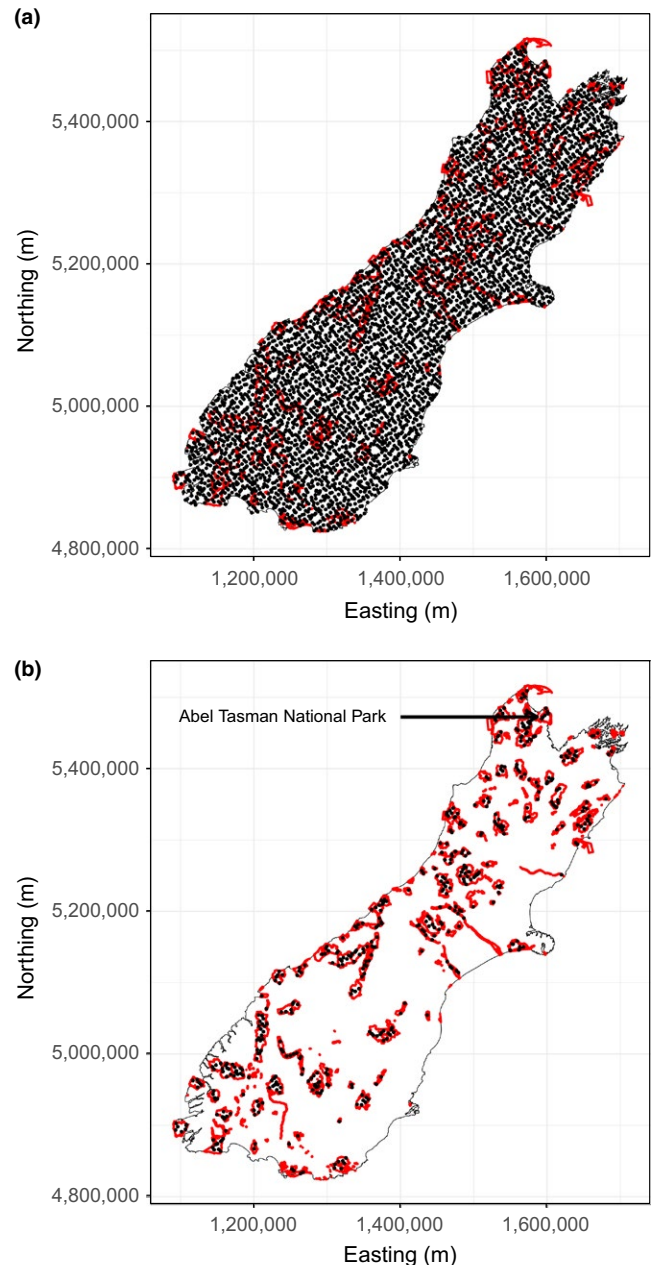
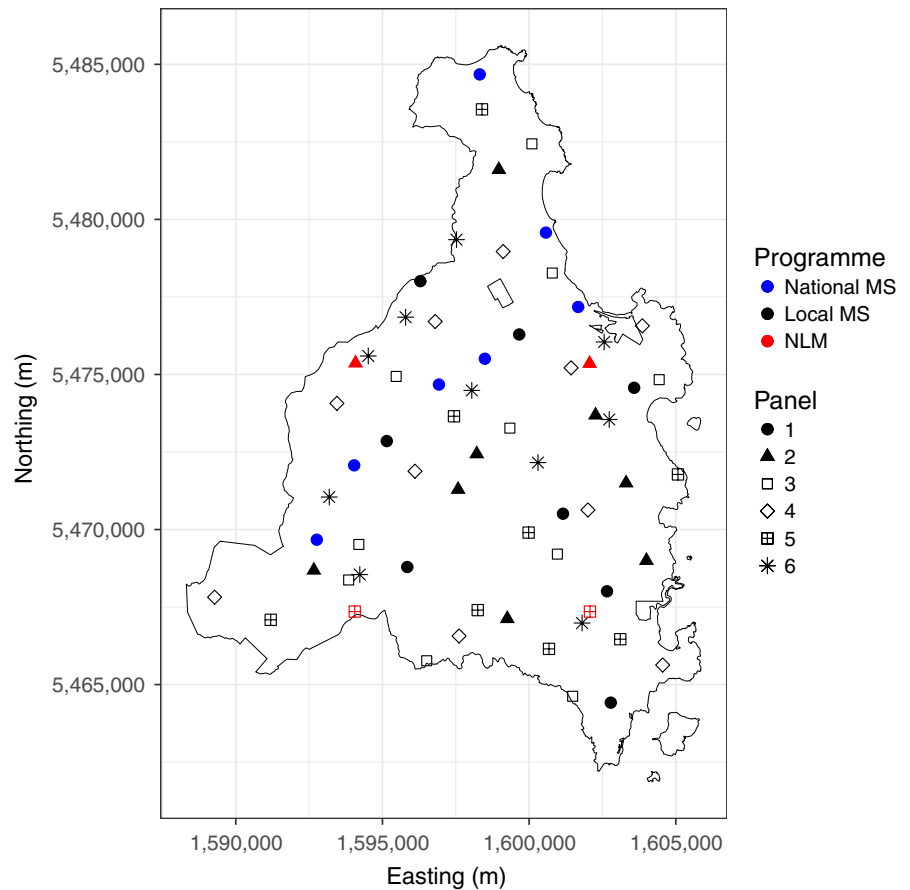


FIGURE 3 South Island of New Zealand. (a) shows the first 5,000 points of the master sample overlaid on red ecosystem management units (EMUs). (b) shows 500 master sample points from (a) that fall within the EMUs in red. Abel Tasman National Park receives seven points which are included as panel 1 points in Figure 4

Department of conservation ecosystem management is focussed on a suite of priority sites known as Ecosystem Management Units (EMUs). To assess the outcome of management interventions across all EMUs, the 8-km grid NLM needs to be intensified with changes to visitation frequency and methods relating to the specific management monitoring objectives. At the same time, there are EMU specific questions to be answered that require another intensification to a regional/local level. For example, one of DOC's EMUs, Abel Tasman National Park (ATNP)

FIGURE 4 An example of bird monitoring in Abel Tasman National Park New Zealand. Panel 1 is measured annually while the other panels are on a 5-year rotation described as $[1-0, (1-4)^5]$. In the first year, panels 1 and 2 would be measured. Blue points are master sample points measured by a national Ecosystem Management Unit monitoring programme that employs the master sample (National MS), see Figure 3. The red points are from the 8-km systematic National Level Monitoring programme (NLM). The black points are the locally augmented sampling units generated from the master sample (Local MS). This design gives excellent spatial coverage over the park each year ($n = 25$) and over a 5-year period ($n = 65$)



(40°56'03.8"S 172°58'19.7"E), is managed in partnership with a philanthropic foundation. The agreement governing this partnership requires the foundation to invest in the recovery of biodiversity in ATNP. Once it is shown that the targets pertaining to increases in the abundance and distribution of bird species and vegetation condition are met, DOC will be responsible for the maintenance of biodiversity in ATNP. Monitoring is required to establish the current state of key indicators in ATNP and then determine when agreed targets have been achieved. A sample design is required which can incorporate the existing NLM monitoring, sample broad scales and enable local intensification of monitoring.

A national sample of NZ using BAS can make up the master sample, with the existing NLM points included. For efficiency and simplicity, each island (North Island, South Island, Stewart Island, Chatham Islands, etc.) will be stratified and have their own bounding box and random seed. The seed chosen for the South Island was $u = (4,887,260, 18,041,662)$ with minimum bounding box in NZ Transverse Mercator 2000 (NZTM2000)

$$[1,089,354, 1,721,164] \times [4,747,979, 5,516,919].$$

These values define the scaling needed to map the random start Halton points to the South Island of NZ. For example, the first point is

$$\begin{aligned} x_1 &= (631,810x_1^{(1)} + 1,089,354, 768,940x_1^{(2)} + 4,747,979) \\ &\approx (1,235,673, 5,075,613). \end{aligned}$$

To sample EMUs in NZ using the master sample, select all sites that fall within EMU polygons. The actual required sample size and visitation frequency should reflect the monitoring objectives and follow a similar process as outlined by Reynolds et al. (2016). See Figure 3 for an example of clipping the master sample on the South Island into the first 500 samples that fall within EMUs.

At ATNP, one of the key targets is focussed on bird abundance and distribution through the park. A sample size of $n = 65$ was chosen based on a precision analysis using *simr* (Green & MacLeod, 2016) and historical bird count data from an existing intensively monitored site which showed that temporal variation was less than spatial. Therefore, 15 points were selected to be measured annually and the other 50 on a rotating 5-year panel $[1-0, (1-4)^5]$ (McDonald, 2003). From the NLM programme mentioned above, there were four legacy samples in the sample frame which would be included in the rotating panel on years corresponding to the years they are to be sampled. If DOC implemented an EMU monitoring programme of 500 augmented points on the South Island, then there would be an additional seven points in ATNP monitored and funded by DOC using the master sample. The EMU points would make up Panel 1 based on the hierarchical order of the master sample. See Figure 4 for the selected points in this monitoring programme. ATNP is an example of localised monitoring using the master sample that can contribute to national estimates of bird abundance and distribution. DOC gets better precision with the increased sampling in ATNP and the philanthropic foundation saves resources by using DOC's national investment in monitoring explicitly in their design.

The master sample above is entirely defined by the seed **u** and the bounding box. Hence, there is no need for a repository to hold the coordinates. Computationally the master sample is easy to run on the fly. Generating 65 points for ATNP in Figure 4 takes ≈ 0.5 s on a desktop computer. See supplementary materials for R script to generate a master sample in NZ.

4 | DISCUSSION

A master sample can be a useful tool to organise environmental monitoring at different spatial scales as previously done using GRTS or RRQRR (Larsen et al., 2008; Theobald, 2016). Using BAS instead of GRTS gives better spatial balance (Robertson et al., 2013) and no need for an oversample. It is also possible to add an extra dimension for unequal probability sampling leading to an overall more flexible design. Not requiring an oversample to create a master sample using BAS means that it will remain relevant to any scale that monitoring takes place no matter how localised. BAS can be used for sampling three-dimensional space (Robertson et al., 2013) which generalises the concepts presented here to work for atmospheric or oceanic monitoring.

Previous master samples rely on large source files for point coordinates. BAS does not because it is deterministic once the random seed is chosen. Generating a BAS master sample in R is computationally quick and easy to program making it possible for a user to run a function in R (see supplementary online material section) to sample a chosen region from a shape file. By making use of the deterministic nature of the Halton sequence and Halton boxes (Robertson et al., 2017) the code can be made computationally efficient for any set of shape files and sample sizes required.

In our experience, any large-scale long-term monitoring will need to incorporate already existing monitoring programmes that are proven effective. This was a requirement in developing a master sample for NZ. We have shown that there is no major issue with incorporating legacy monitoring into the design but recommend that the sample is rigorously vetted to ensure no known biases are included, for example, the legacy monitoring is a judgement sample. Using panel designs can help incorporate the already existing visitation schedule of the legacy units into an efficient monitoring design. By generating the master sample independently of the legacy monitoring it is possible that a legacy unit and master sample unit could be close in space, in this case both units still need to be measured.

The master sample helps coordinate the points sampled for environmental surveys. Every survey at the local and national level should still go through rigorous design. This means defining the objectives of monitoring clearly and the methods to use so that they are consistent with standard methodology as required by the objective. By following the steps outlined in Reynolds et al. (2016) and using the master sample for point generation, we believe that the monitoring programmes undertaken at all levels will have improved efficiency and contribute to the overall knowledge of the population of interest.

ACKNOWLEDGEMENTS

This research was funded by the Department of Conservation in New Zealand. We thank Elaine Wright for her support and Tony Olsen, Scott Foster and Amy Hawcroft for their advice.

AUTHORS' CONTRIBUTIONS

All authors had significant contributions to this research. The paper was initiated by a requirement from the Department of Conservation (DOC) for coordinated monitoring of ecosystems at the national level. This work fell directly to O.G. and P.v.D.-B. who collaborated with B.R. We were funded as salary through our roles at DOC and the University of Canterbury.

DATA ACCESSIBILITY

We present a maintained version of the R code to generate a master sample in New Zealand, including a shape file example on GitHub for access to readers familiar with R <https://doi.org/10.5281/zenodo.1193953>.

ORCID

Paul van Dam-Bates  <http://orcid.org/0000-0002-8094-1518>

REFERENCES

- Aune-Lundberg, L., & Strand, G.H. (2014). Comparison of variance estimation methods for use with two-dimensional systematic sampling of land use/land cover data. *Environmental Modelling & Software*, 61, 87–97. <https://doi.org/10.1016/j.envsoft.2014.07.001>
- Buckland, S., Magurran, A., Green, R., & Fewster, R. (2005). Monitoring change in biodiversity through composite indices. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360, 243–254. <https://doi.org/10.1098/rstb.2004.1589>
- Cochran, W. G. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. *The Annals of Mathematical Statistics*, 17, 164–177. <https://doi.org/10.1214/aoms/1177730978>
- Collier, K. J., & Olsen, A. R. (2013). Monitoring network-design influence on assessment of ecological condition in Wadeable streams. *Marine and Freshwater Research*, 64, 146–156. <https://doi.org/10.1071/mf12267>
- Coomes, D. A., Allen, R. B., Scott, N. A., Goulding, C., & Beets, P. (2002). Designing systems to monitor carbon stocks in forests and shrublands. *Forest Ecology and Management*, 164, 89–108. [https://doi.org/10.1016/S0378-1127\(01\)00592-8](https://doi.org/10.1016/S0378-1127(01)00592-8)
- Department of Conservation (2016). Monitoring species and habitats in terrestrial, freshwater and marine environments. Retrieved from <http://www.doc.govt.nz/our-work/biodiversity-inventory-and-monitoring/> (last accessed: 29 September 2016).
- Dixon, P. M., Olsen, A. R., & Kahn, B. M. (1998). Measuring trends in ecological resources. *Ecological Applications*, 8, 225–227. [https://doi.org/10.1890/1051-0761\(1998\)008\[0225:mtier\]2.0.co;2](https://doi.org/10.1890/1051-0761(1998)008[0225:mtier]2.0.co;2)
- Fancy, S., Gross, J., & Carter, S. (2009). Monitoring the condition of natural resources in US National Parks. *Environmental Monitoring and Assessment*, 151, 161–174. <https://doi.org/10.1007/s10661-008-0257-y>

- Field, S. A., O'Connor, P. J., Tyre, A. J., & Possingham, H. P. (2007). Making monitoring meaningful. *Austral Ecology*, 32, 485–491.
- Foster, S. D. (2016). *MBHdesign: Spatial designs for ecological and environmental surveys*. R package version 1.0.61. Retrieved from <https://CRAN.R-project.org/package=MBHdesign>
- Foster, S. D., Hosack, G. R., Lawrence, E., Przeslawski, R., Hedge, P., Caley, M. J., ... Hayes, K. R. (2017). Spatially-balanced designs that incorporate legacy sites. *Methods in Ecology and Evolution*, 8, 1433–1442. <https://doi.org/10.1111/2041-210x.12782>
- Grafström, A., Lundström, N. L. P., & Schelin, L. (2012). Spatially balanced sampling through the pivotal method. *Biometrics*, 68, 514–520. <https://doi.org/10.1111/j.1541-0420.2011.01699.x>
- Grafström, A., & Lundström, N. L. (2013). Why well spread probability samples are balanced. *Open Journal of Statistics*, 3, 36–41. <https://doi.org/10.4236/ojs.2013.31005>
- Grafström, A., Saarela, S., & Ene, L. T. (2014). Efficient sampling strategies for forest inventories by spreading the sample in auxiliary space. *Canadian Journal of Forest Research*, 44, 1156–1164. <https://doi.org/10.1139/cjfr-2014-0202>
- Green, P., & MacLeod, C. J. (2016). simr: An R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, 7, 493–498. <https://doi.org/10.1111/2041-210x.12504>
- Halton, J. H. (1960). On the efficiency of certain Quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2, 84–90. <https://doi.org/10.1007/bf01386213>
- Kincaid, T. M., & Olsen, A. R. (2016). *spsurvey: Spatial survey design and analysis*. R package version 3.3. Retrieved from <https://cran.r-project.org/package=MBHdesign>
- King, A. J. (1945). The master sample of agriculture: I. Development and use. *Journal of the American Statistical Association*, 40, 38–45. <https://doi.org/10.2307/2280452>
- Larsen, D. P., Olsen, A. R., & Stevens, D. (2008). Using a master sample to integrate stream monitoring programs. *Journal of Agricultural, Biological, and Environmental Statistics*, 13, 243–254. <https://doi.org/10.1198/108571108x336593>
- Legg, C. J., & Nagy, L. (2006). Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management*, 78, 194–199. <https://doi.org/10.1016/j.jenvman.2005.04.016>
- Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., ... Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: Assessing change in ecological communities through time. *Trends in Ecology & Evolution*, 25, 574–582. <https://doi.org/10.1016/j.tree.2010.06.016>
- McDonald, T. L. (2003). Review of environmental monitoring methods: Survey designs. *Environmental Monitoring and Assessment*, 85, 277–292. <https://doi.org/10.1023/A:1023954311636>
- Nichols, J. D., & Williams, B. K. (2006). Monitoring for conservation. *Trends in Ecology & Evolution*, 21, 668–673. <https://doi.org/10.1016/j.tree.2006.08.007>
- Noss, R. F. (1999). Assessing and monitoring forest biodiversity: A suggested framework and indicators. *Forest Ecology and Management*, 115, 135–146. [https://doi.org/10.1016/S0378-1127\(98\)00394-6](https://doi.org/10.1016/S0378-1127(98)00394-6)
- Pereira, H. M., & Cooper, H. D. (2006). Towards the global monitoring of biodiversity change. *Trends in Ecology & Evolution*, 21, 123–129. <https://doi.org/10.1016/j.tree.2005.10.015>
- Peterson, S. A., Urquhart, N. S., & Welch, E. B. (1999). Sample representativeness: A must for reliable regional lake condition estimates. *Environmental Science & Technology*, 33, 1559–1565. <https://doi.org/10.1021/es980711l>
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Reynolds, J. H., Knutson, M. G., Newman, K. B., Silverman, E. D., & Thompson, W. L. (2016). A road map for designing and implementing a biological monitoring program. *Environmental Monitoring and Assessment*, 188, 1–25. <https://doi.org/10.1007/s10661-016-5397-x>
- Robertson, B. L., Brown, J. A., McDonald, T. L., & Jaksons, P. (2013). BAS: Balanced acceptance sampling of natural resources. *Biometrics*, 69, 776–784. <https://doi.org/10.1111/biom.12059>
- Robertson, B. L., McDonald, T. L., Price, C. J., & Brown, J. A. (2017). A modification of balanced acceptance sampling. *Statistics & Probability Letters*, 129, 107–112. <https://doi.org/10.1016/j.spl.2017.05.004>
- Skalski, J. R. (1990). A design for long-term status and trends monitoring. *Journal of Environmental Management*, 30, 139–144. [https://doi.org/10.1016/0301-4797\(90\)90011-k](https://doi.org/10.1016/0301-4797(90)90011-k)
- Stevens, D. (1997). Variable density grid-based sampling designs for continuous spatial populations. *Environmetrics*, 8, 167–195. [https://doi.org/10.1002/\(sici\)1099-095x\(199705\)8:3<167::aid-env239>3.0.co;2-d](https://doi.org/10.1002/(sici)1099-095x(199705)8:3<167::aid-env239>3.0.co;2-d)
- Stevens, D., & Olsen, A. R. (2003). Variance estimation for spatially balanced samples of environmental resources. *Environmetrics*, 14, 593–610. <https://doi.org/10.1002/env.606>
- Stevens, D., & Olsen, A. R. (2004). Spatially balanced sampling of natural resources. *Journal of the American Statistical Association*, 99, 262–278. <https://doi.org/10.1198/016214504000000250>
- Stevens, D., & Olsen, A. R. (1999). Spatially restricted surveys over time for aquatic resources. *Journal of Agricultural, Biological, and Environmental Statistics*, 4, 415–428.
- Strand, G. H. (2017). A study of variance estimation methods for systematic spatial sampling. *Spatial Statistics*, 21, 226–240. <https://doi.org/10.1016/j.spasta.2017.06.008>
- Theobald, D. M. (2016). A general-purpose spatial survey design for collaborative science and monitoring of global environmental change: The global grid. *Remote Sensing*, 8, 813. <https://doi.org/10.3390/rs8100813>
- Theobald, D. M., Stevens, D. L., White, D., Urquhart, N. S., Olsen, A. R., & Norman, J. B. (2007). Using GIS to generate spatially balanced random survey designs for natural resource applications. *Environmental Management*, 40, 134–146. <https://doi.org/10.1007/s00267-005-0199-x>
- Thompson, W. L., Miller, A. E., Mortenson, D. C., & Woodward, A. (2011). Developing effective sampling designs for monitoring natural resources in Alaskan National Parks: An example using simulations and vegetation data. *Biological Conservation*, 144, 1270–1277. <https://doi.org/10.1016/j.biocon.2010.09.032>
- Urquhart, N. S., & Kincaid, T. M. (1999). Designs for detecting trend from repeated surveys of ecological resources. *Journal of Agricultural, Biological, and Environmental Statistics*, 4, 404–414. <https://doi.org/10.2307/1400498>
- Wang, X., & Hickernell, F. J. (2000). Randomized halton sequences. *Mathematical and Computer Modelling*, 32, 887–899. [https://doi.org/10.1016/S0895-7177\(00\)00178-3](https://doi.org/10.1016/S0895-7177(00)00178-3)
- Yoccoz, N. G., Nichols, J. D., & Boulinier, T. (2001). Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution*, 16, 446–453. [https://doi.org/10.1016/S0169-5347\(01\)02205-4](https://doi.org/10.1016/S0169-5347(01)02205-4)

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: van Dam-Bates P, Gansell O, Robertson B. Using balanced acceptance sampling as a master sample for environmental surveys. *Methods Ecol Evol*. 2018;00:1–9. <https://doi.org/10.1111/2041-210X.13003>